

AD-A132 277

A GENERAL FRAMEWORK FOR LEARNING CURVE RELIABILITY
GROWTH MODELS(U) CALIFORNIA UNIV BERKELEY OPERATIONS
RESEARCH CENTER W S JEWELL APR 83 ORC-79-11

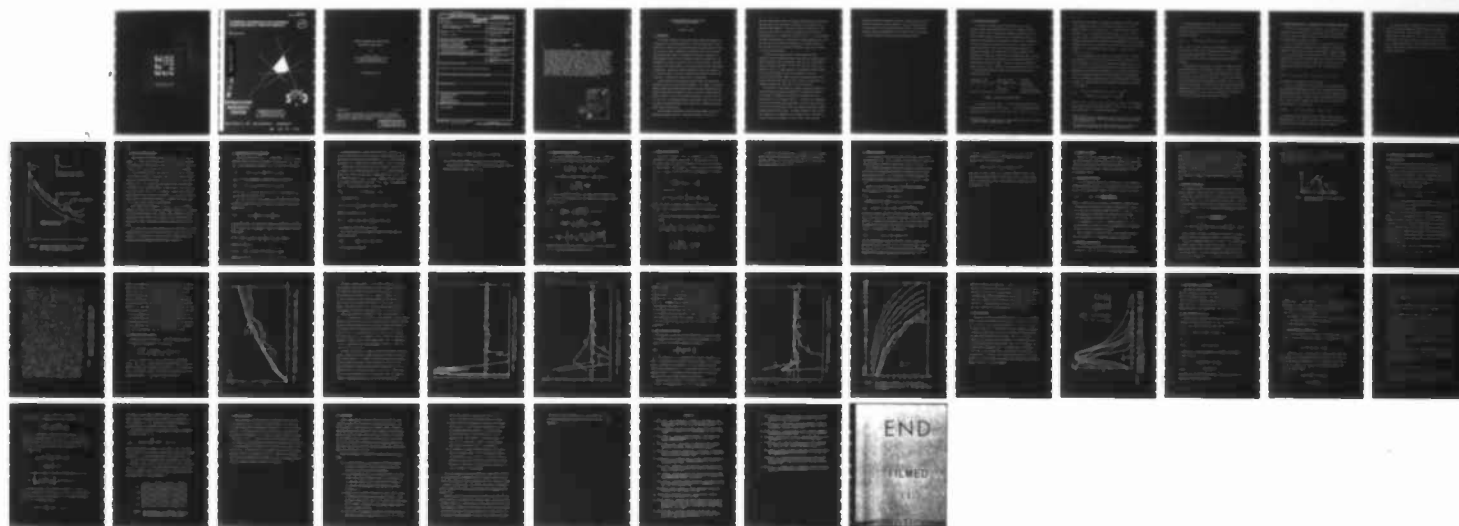
1/1

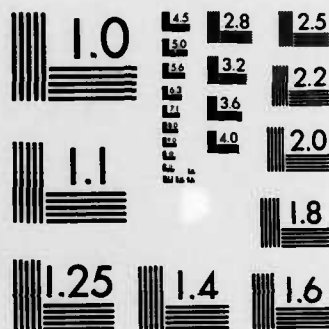
UNCLASSIFIED

AFOSR-81-0122

F/G 12/1

NL





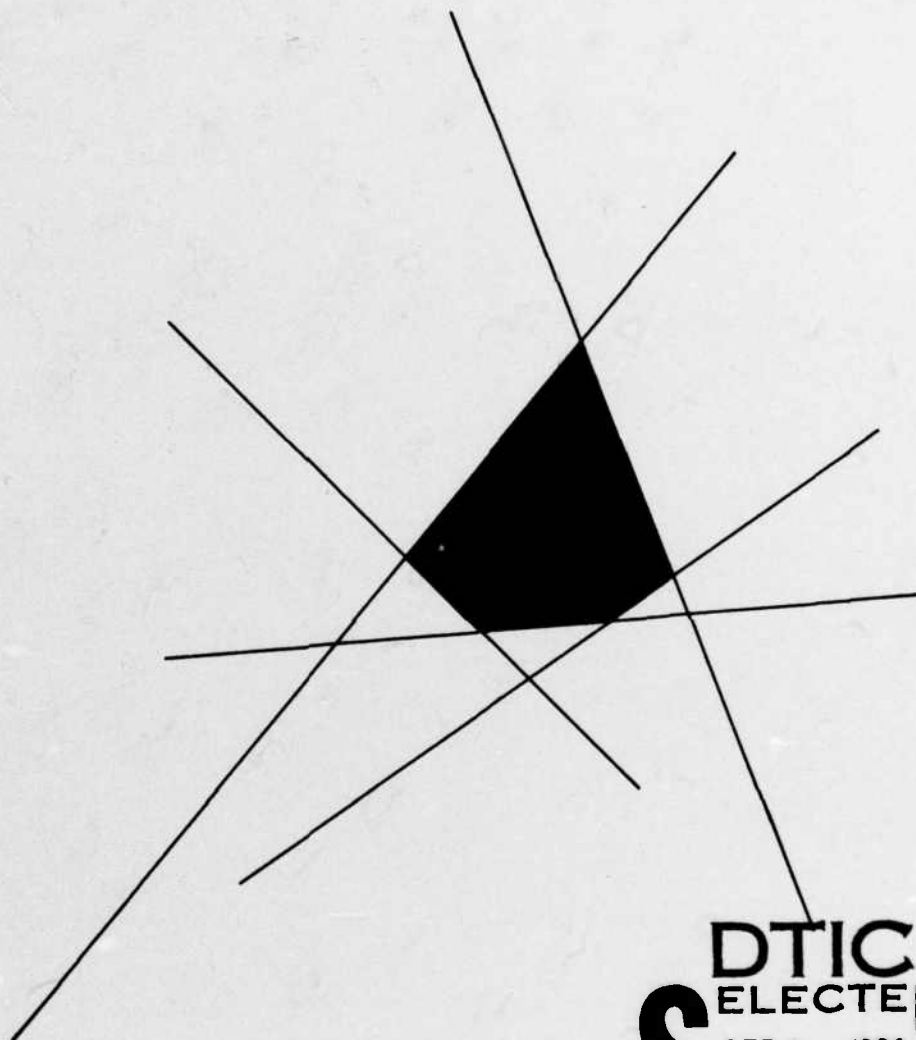
A GENERAL FRAMEWORK FOR LEARNING CURVE RELIABILITY GROWTH MODELS

12

by
WILLIAM S. JEWELL

AD A132277

DTIC FILE COPY



DTIC
ELECTE
SEP 9 1983
S B D

OPERATIONS
RESEARCH
CENTER

DISTRIBUTION STATEMENT A

Approved for public release
Distribution Unlimited

UNIVERSITY OF CALIFORNIA • BERKELEY

83 09 07 008

A GENERAL FRAMEWORK FOR LEARNING CURVE
RELIABILITY GROWTH MODELS

by

William S. Jewell
Department of Industrial Engineering
and Operations Research
University of California, Berkeley

(REVISED APRIL 1983)

OCTOBER 1979

ORC 79-11

This research was supported by the Air Force Office of Scientific Research (AFSC), USAF, under Grant AFOSR-81-0122 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ORC 79-11	2. GOVT ACCESSION NO. A132277	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A GENERAL FRAMEWORK FOR LEARNING CURVE RELIABILITY GROWTH MODELS		5. TYPE OF REPORT & PERIOD COVERED Research Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William S. Jewell		8. CONTRACT OR GRANT NUMBER(s) AFOSR-81-0122
9. PERFORMING ORGANIZATION NAME AND ADDRESS Operations Research Center University of California Berkeley, California 94720		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS United States Air Force Air Force Office of Scientific Research Bolling Air Force Base, D.C. 20332		12. REPORT DATE October 1979 (Revised April 1983)
		13. NUMBER OF PAGES 47
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Reliability Growth Learning Curves Failure Rate Maximum Likelihood Bayesian Prediction		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) (SEE ABSTRACT)		

DD FORM 1473
1 JAN 73

EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT

In reliability growth models, systems undergo an improvement in performance during prototype testing, as design changes are made, and operating procedures and environment are modified. In the learning-curve models, this improvement occurs continuously over time, and there is great interest in predicting the ultimate performance of the system, using only the epochs of the failures which occur early in the testing program. This paper constructs a general framework in which to analyze this problem, including as special cases many different model variations that have previously been analyzed. Numerical trials indicate the difficulty of using classical procedures to estimate ultimate performance; the maximum likelihood estimator is unstable for small testing intervals with a small number of systems on test, and is even inconsistent for a large number of systems. Bayesian procedures are recommended for implementation, as they can use the data from any testing protocol.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	



A GENERAL FRAMEWORK FOR LEARNING CURVE RELIABILITY GROWTH MODELS

by

William S. Jewell

1. INTRODUCTION

When complex systems are first produced and undergo development and testing prior to actual operational use, a *learning curve* effect usually takes place, in which the initially high failure rate diminishes over time as the various causes of breakdown are identified and corrected through engineering design changes, environmental modification, operating procedure changes, and so on. This "shakedown" process is essential to the development of a robust system that can be routinely produced, installed, and operated with high reliability under long-duration, adverse-environment circumstances. Because of the high development costs of such systems, it is clearly important to be able to estimate this *reliability growth*, and, especially, to estimate the ultimate level of performance.

A large variety of special-purpose growth models have already been proposed; [7] and [10] are convenient summaries. For example, one group of models focuses upon the *defect-identification* process, assuming that each system failure gives the engineer a chance to rectify a newly-discovered cause of failure. Most of these models are Markovian in nature; see, e.g., [2], [3], [6]. The primary drawback of these models seems to be the need to specify an appropriate defect state space, and to estimate the related defect removal probabilities.

A second method of modelling reliability growth might be called the *structural parameter* approach. Learning is assumed to occur only in

discrete, well-defined stages, with unknown failure parameters associated with each stage; only the number of failures in each stage is used to estimate these parameters. If growth is assumed to take place at each stage, then this leads to a monotone estimation problem in which the many parameters are, by hypothesis, ordered [2], [21]. It is also possible to weaken this assumption to stochastically ordered parameters [15], [20]. The most serious criticism of this approach to reliability growth is that it gives very little predictive capability, since future learning is only loosely constrained.

A third group of models requires strong assumptions about the general form of reliability growth, based upon past experiences with similar processes; then one estimates a relatively few free parameters using only failure epoch data. We will call this the *learning curve* approach, since it is closely related to deterministic models of production learning, developed by industrial engineers in the 50's. A large literature exists, mostly focused around special growth curve forms, such as the *Duane model* [7], [10]. Many of the recent software reliability models [11], [18], [19] are also of this type.

The objective of this paper is to present a general framework for constructing and estimating the parameters of learning-curve models, in particular, two important variations of the general form, the "as operated," and "as produced" models. It is shown that, independent of the form of learning deemed appropriate, the type of model selected as well as the testing protocol used, the parameter estimation problems are mathematically very similar. Preliminary numerical trials indicate that classical point estimators of the parameters may be misleading as to their accuracy; accordingly, we shall outline several possible

Bayesian approaches to parameter estimation. Detailed results from our numerical simulation are then presented. Our main conclusion is that most learning-curve testing protocols provide little of the parameter precision needed by manufacturers to make "early returns" predictions of ultimate reliability using traditional estimation methods.

2. FAILURE MODEL FORMULATION

In developing a general model of reliability growth, we must consider carefully the assumptions made about the underlying failure process. First of all, by ignoring any specific information about the *type* of failure (except to assure that this information is used in a positive, predictable way by the engineers to modify future performance), we restrict ourselves to observing *point processes*. Figure 1 shows a single system on test, with various failure events ("glitches") indicated by asterisks. Our objective is to develop a stochastic law governing the occurrence of these events; in reliability growth, we expect these events to occur less frequently, or to be more widely spaced, as time evolves from the test origin. (Section 7 discusses results with several systems on test.) Because of the concerns expressed in [14], [16], a failure-rate-oriented model appears to be more appropriate than an interval-oriented one.

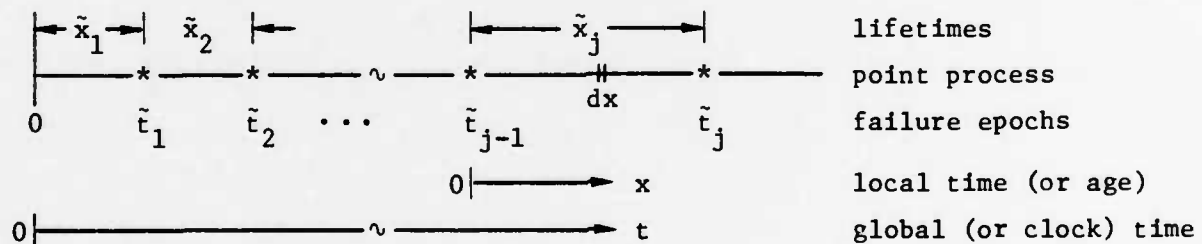


FIGURE 1. Notation with Single System.

To fix notation,^{*} let $\tilde{x}_1, \tilde{x}_2, \dots$ be the successive random *lifetimes* (intervals between failures), and $\tilde{t}_1, \tilde{t}_2, \dots$ be the corresponding failure event *epochs*; clearly $\tilde{t}_0 = 0$, and $\tilde{t}_j = \tilde{x}_1 + \tilde{x}_2 + \dots + \tilde{x}_j$ ($j = 1, 2, \dots$).

^{*}A random variable is indicated with a tilde; its actual value in some realization is the same variable without a tilde.

Suppose we are formulating the stochastic law governing the j^{th} lifetime. It is convenient to refer to a small interval of time, dx , which is measured either in *local time* (or *age*), x , since the actual last failure epoch, t_{j-1} , or in *global* (or *clock*) time, $t = t_{j-1} + x$, from the test origin.* A failure-rate law for the j^{th} interval would then be completely determined if we could specify the probability of the next failure epoch occurring in dx as a function of the *entire past history* of the point process, i.e., as a function of, say, $(x_1, x_2, \dots, x_{j-1}; x)$, or, of $(t_1, t_2, \dots, t_{j-1}; t)$.

Our first assumption will be that *the starting epoch of an interval is a surrogate for all previous failure history of the system*; i.e., for the j^{th} interval, \tilde{x}_j and \tilde{t}_j are statistically dependent only on $\tilde{t}_{j-1} = t_{j-1}$, and not on the other failure epochs (and not upon the index j). Physically, this means that the global time t_{j-1} describes sufficiently all of the engineering changes, the characteristics of the replacement items, the procedural changes that have been incorporated into the system to date, etc. Mathematically, we are specifying a failure rate function, f , such that:**

$$(2.1) \quad \begin{aligned} \Pr \{ \tilde{x}_j \in (x, x+dx) \mid \tilde{x}_j > x ; x_1, x_2, \dots, x_{j-1}; t \} \\ = f(x; t \mid t_{j-1} = x_1 + x_2 + \dots + x_{j-1}) dx \end{aligned}$$

for $(j = 1, 2, \dots)$, $(0 \leq x \leq x_j)$, and $(t_{j-1} < t \leq t_j)$. For simplicity in the sequel, we will move freely between t and $x = t - t_{j-1}$, whenever the interval index j is clear from the context.

* Time and age can also be measured in *production time* units, or in actual *system operational time*. This raises certain measurement difficulties, but no conceptual ones.

** We emphasize the case where the random variables are continuous, but discrete cases follow directly, with only minor modifications.

The second assumption we shall make is that *the age-dependent and time-dependent failure mechanisms of the system are statistically independent*, e.g., that the failure rate function is the sum of two terms, so that for interval j ,

$$(2.2) \quad f(x; t \mid t_{j-1}) = \phi g(t \mid t_{j-1}) + \omega h(x) .$$

Here ϕ and ω are two parameters to be estimated, and g and h are two *prototypical growth forms*, assumed known from previous experience. h is the usual age-dependent *hazard function*, while g is the *learning-curve* itself, which is decreasing in some given manner as t increases.

This assumption is reasonable if we think of a single unreliable item being replaced, with the "wear-out" hazard h being a property of this item, while g reflects independent improvement of the rest of the system or the operational environment. If the system consists of many unreliable terms with exponential lifetimes ($h(x) = 1$), then this model also applies, if the reliability growth provides independent environmental improvement, or if it rectifies or removes certain of the unreliable elements [5], [18]. Further development or validation of (2.2) under more complex system assumptions is needed.

3. ALTERNATE FORMULATIONS: "AS OPERATED" AND "AS PRODUCED" GROWTH MODELS

The choice of the learning curve function, g , depends upon further assumptions about the nature of the improvements made upon the system over time. We shall consider two special cases which we believe bracket the possibilities to be found in practice.

Type I ("as operated") models refer to those situations in which the reliability growth is either due to improvements in environmental conditions (cooling, shock insulation, power supplies, support equipment etc.), continuing changes in operational procedures, or when any design improvements on current production models can be retrofitted on existing systems in the field or on test. In this case, we assume that a one-dimensional prototypical learning curve $g(t)$ is known, and that in (2.2):

$$(3.1I) \quad g(t \mid t_{j-1}) = g(t), \quad (t_{j-1} < t \leq t_j = t_{j-1} + x_j).$$

Most models in the literature are of this type.

Type II ("as produced") models will refer to the other extreme in which reliability improvements are built into each system only during the production-installation process, and thus do not affect equipment already in the field or on test. As above, there is an underlying learning curve $g(t)$, but for the j^{th} interval, only the value of the function at the starting epoch is of interest, and

$$(3.1II) \quad g(t \mid t_{j-1}) = g(t_{j-1}) \quad (t_{j-1} < t \leq t_j)$$

is a constant over this lifetime interval. Models of this type were introduced in [4], [5], [18]. Note that a Duane growth curve (see below) has some technical difficulty in the first interval of model II.

One can easily imagine combinations of these two types of models, as when retrofitting engineering changes gives less improvement when applied to used items. However, Models I and II are clearly good approximations to each other when g is a slowly varying function, as compared to the random lifetimes, so that, in many situations, we may not expect these models to give dramatically different results. Figure 2 shows how the actual failure curves for the two different models would be constructed.

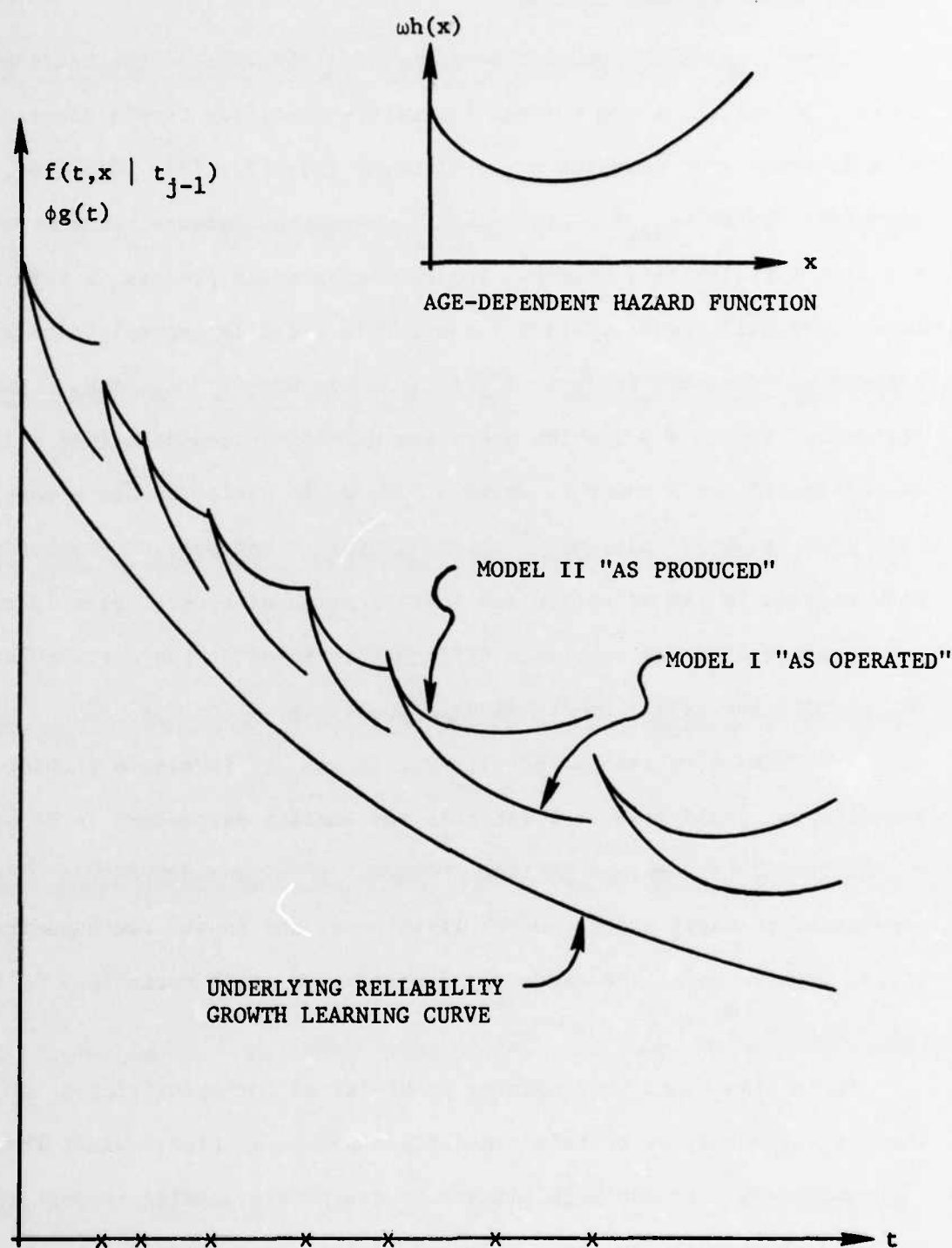


FIGURE 2. Components of failure rates, ωh and ϕg , and resulting total failure rates, f , of Models I and II for a particular realization.

4. CHOICE OF GROWTH FUNCTIONS

There remains the problem of specifying the form of the learning curve, g . Much of the current literature specifies Crow's adaptation of a deterministic learning curve of Duane [1], [7], [8], [9], [10], [17], in which $f(x, t \mid t_{j-1}) = g(t) = kt^{\gamma-1}$, with the internal parameter $\gamma < 1$ for reliability growth. The resulting point process is referred to as a Weibull process. In our view, this model is incomplete because there is no residual failure effect ($\omega = 0$) at $t \rightarrow \infty$. Also, this curve "blows up" for $t = 0$, which makes the normalizations described below in Sections 5 and 9 somewhat awkward. We would prefer to use a more tractable form, such as $g(t) = e^{-\gamma t}$, for example. However, much practical work remains in the selection and justification of a particular learning curve, since it seems extremely difficult to identify the correct form of g with the usual amounts of data available.

The "hidden" shape parameter, γ , in any g is also a problem. Normally, one would like to leave this and similar parameters to be estimated by the procedure described below. However, these parameters are quite non-linear in their effect on the likelihood, and in the few numerical trials we have made, are extremely sensitive to small variations in the data.

It is also clear that nothing prohibits us for specifying a g which changes value only at certain pre-defined points in time, rather than continuously. But perhaps such situations are better handled through the discrete structural parameter models described in Section 1.

5. DISTRIBUTIONS AND DATA LIKELIHOODS

The random lifetimes, $\tilde{x}_1, \tilde{x}_2, \dots$, generated by (2.2) are, in general, dependent and not identically distributed. However, given t_{j-1} (and ϕ and ω) the conditional distribution of \tilde{x}_j is found from the relation

$$(5.1) \quad f(t, x \mid t_{j-1}) = -\frac{d}{dx} \ln P^c(x \mid t_{j-1}; \phi, \omega),$$

where P^c is the complementary distribution function defined by

$$(5.2) \quad P^c(x \mid t_{j-1}; \phi, \omega) = \Pr \{ \tilde{x}_j > x \mid t_{j-1}; \phi, \omega \},$$

for $(j = 1, 2, \dots)$ $(x \geq 0)$ and $(t = t_{j-1} + x)$. The density of \tilde{x}_j is just the negative derivative of P^c .

Let $G(t)$ and $H(x)$ be the integrals of the growth and age hazard components, measured from the test origin and the current interval origin, respectively:

$$(5.3) \quad G(t) = \int_0^t g(u) du; \quad H(x) = \int_0^x h(w) dw.$$

(As described above, this raises minor technical difficulties in the Duane model in which $G(0)$ is undefined.) Then it is straightforward that for *model type I*:

$$(5.4I) \quad P^c(x \mid t_{j-1}; \phi, \omega) = \exp \{ -\phi[G(t) - G(t_{j-1})] - \omega H(x) \},$$

whereas for *type II*:

$$(5.4II) \quad P^c(x \mid t_{j-1}; \phi, \omega) = \exp \{ -\phi x g(t_{j-1}) - \omega H(x) \},$$

remembering always that $t = t_{j-1} + x$ in the current interval j .

Having developed the law for each interval of our point process, we can now consider specific testing protocols, calculate appropriate data likelihoods, and proceed with the estimation of ϕ , ω , and any hidden parameters in g . Remember that we have assumed g and h are known.

Suppose there is one system on test, as in Figure 1, and the test is stopped at some predetermined global time $t = T$ ("time-truncated"). During $(0, T]$, there will have been a random number $\tilde{n}(T) = \sup \{n \mid \tilde{t}_n \leq T\}$ of complete lifetimes, $x_1, x_2, \dots, x_{\tilde{n}(T)}$, together with the $(\tilde{n}(T)+1)^{\text{st}}$ "in progress" lifetime with current age $x = T - t_{\tilde{n}(T)}$. The appropriate data, \mathcal{D} , from our point process will then be:

$$(5.5) \quad \mathcal{D} = \{n; T; x_1, x_2, \dots, x_n\}.$$

The likelihood is:

$$(5.6) \quad L(\phi, \omega \mid \mathcal{D}) = \prod_{j=1}^n [\phi g_j(\mathcal{D}) + \omega h_j(\mathcal{D})] \exp \{-\phi G(\mathcal{D}) - \omega H(\mathcal{D})\},$$

where for both models I and II:

$$(5.7) \quad h_j(\mathcal{D}) = h(x_j); \quad H(\mathcal{D}) = \sum_{j=1}^n H(x_j) + H(T - t_n).$$

H is called the "total-Q-on-test" in [12].

The learning curve data components depend on the model type; for growth "as operated":

$$(5.8I) \quad g_j(\mathcal{D}) = g(t_j); \quad G(\mathcal{D}) = G(T);$$

while for growth "as produced":

$$(5.8II) \quad g_j(\mathcal{D}) = g(t_{j-1}) ; G(\mathcal{D}) = \sum_{j=1}^n x_j g(t_{j-1}) + (T - t_n) g(t_n) .$$

It is clear that these components are quite close for slowly varying g .

If observations are stopped at the n^{th} failure epoch ("failure truncated"), the above formula still apply, with $T = t_n$.

6. MAXIMUM-LIKELIHOOD ESTIMATION

The classical maximum-likelihood estimates $\hat{\phi}$ and $\hat{\omega}$ follow directly from (5.6); we get an implicit solution through the two equations:

$$(6.1) \quad \sum_{j=1}^n \frac{h_j}{\hat{\phi}g_j + \hat{\omega}h_j} = H ; \quad \sum_{j=1}^n \frac{g_j}{\hat{\phi}g_j + \hat{\omega}h_j} = G .$$

If g also contains an internal parameter γ , then we have at $\gamma = \hat{\gamma}$

$$(6.2) \quad \sum_{j=1}^n \frac{\frac{\partial g_j(\gamma)}{\partial \gamma}}{\hat{\phi}g_j + \hat{\omega}h_j} = \frac{\partial G(\gamma)}{\partial \gamma} ,$$

in an obvious extension of notation. For computational convenience, we record also the second derivatives of \ln -likelihood at the maximum likelihood point $(\hat{\phi}, \hat{\omega}, \hat{\gamma})$:

$$(6.3) \quad \frac{\partial^2 \ln L}{\partial \phi^k \partial \omega^{2-k}} = - \sum_{j=1}^n \frac{(g_j)^k (h_j)^{2-k}}{(\hat{\phi}g_j + \hat{\omega}h_j)^2} ; \quad (k = 0, 1, 2)$$

$$(6.4) \quad \frac{1}{\hat{\omega}} \frac{\partial^2 \ln L}{\partial \phi \partial \gamma} = \sum_{j=1}^n \frac{h_j \frac{\partial g_j}{\partial \gamma}}{(\hat{\phi}g_j + \hat{\omega}h_j)^2} = - \frac{1}{\hat{\phi}} \frac{\partial^2 \ln L}{\partial \omega \partial \gamma} ;$$

$$(6.5) \quad \frac{\partial^2 \ln L}{\partial \gamma^2} = \hat{\phi} \left[- \frac{\partial^2 G}{\partial \gamma^2} + \sum_{j=1}^n \frac{(\hat{\phi}g_j + \hat{\omega}h_j) \frac{\partial^2 g_j}{\partial \gamma^2} - \hat{\phi} \left(\frac{\partial g_j}{\partial \gamma} \right)^2}{(\hat{\phi}g_j + \hat{\omega}h_j)^2} \right] .$$

Because these derivatives are evaluated at $[\hat{\phi}, \hat{\omega}, \hat{\gamma}]$, they are also equal to L^{-1} times the corresponding derivatives of L .

7. OTHER TESTING PROTOCOLS

The results of Section 6 can be easily adapted to other testing arrangements. For example, suppose that S systems were put on test at the same global time origin, and operated for the same total interval T . Then, system i would report n_i failures, with completed lifetimes $x_{i1}, x_{i2}, \dots, x_{in_i}$, and the total data set corresponding to (5.5) would be:

$$(7.1) \quad \mathcal{D} = \bigcup_{i=1}^S \left\{ n_i; T; x_{i1}, x_{i2}, \dots, x_{in_i} \right\},$$

with (5.7) replaced by:

$$(7.2) \quad h_{ij}(\mathcal{D}) = h(x_{ij}) ; H_i(\mathcal{D}) = \sum_{j=1}^{n_i} H(x_{ij}) + H(T - t_{in_i}) .$$

$g_{ij}(\mathcal{D})$ and $G_i(\mathcal{D})$ are defined analogously to (5.8I) or (5.8II); then the total $L(\mathcal{D})$ is the product of terms like (5.6) for each system $i = 1, 2, \dots, S$.

It follows that the maximum likelihood estimator formulae (6.1) and (6.2) become:

$$(7.3) \quad \sum_{i=1}^S \sum_{j=1}^{n_i} \frac{h_{ij}}{\hat{\phi}g_{ij} + \hat{\omega}h_{ij}} = \sum_{i=1}^S H_i ; \quad \sum_{i=1}^S \sum_{j=1}^{n_i} \frac{g_{ij}}{\hat{\phi}g_{ij} + \hat{\omega}h_{ij}} = \sum_{i=1}^S G_i ;$$

and

$$(7.4) \quad \sum_{i=1}^S \sum_{j=1}^{n_i} \frac{\frac{\partial g_{ij}(\gamma)}{\partial \gamma}}{\hat{\phi}g_{ij} + \hat{\omega}h_{ij}} = \sum_{i=1}^S \frac{\partial G_i(\gamma)}{\partial \gamma} .$$

A variety of other testing protocols can easily be incorporated into similar formulae, including different test origins and terminations, missing intervals, etc.; the stopping time(s) may also depend upon the data outcome (see, e.g., [12], [13]).

8. CONSTANT AGE HAZARD

It is interesting to note that only the ratios g_i/h_i entered into the computations (6.1) - (6.5); in effect, one could switch to an *operational time* and analyze an equivalent model in which $h(x) \equiv 1$ for all x .

This constant age-hazard is important in ordinary reliability applications, where it corresponds to exponential lifetimes and the stationary Poisson failure process. It is of interest to see what happens in the corresponding reliability growth model, with $h_i = 1$ for all i , and $H = T$.

Models of type I then give rise to time-varying (inhomogeneous) Poisson processes, with intensity parameter

$$(8.1) \quad \lambda(t) = \phi g(t) + \omega.$$

The distribution of the number of failures in $(0, T]$ is Poisson:

$$(8.2) \quad \Pr \{ \tilde{n}(T) = n \mid \phi, \omega \} = [\phi G(T) + \omega T]^n e^{-[\phi G(T) + \omega T]} / n!$$

from which we can get the law of n^{th} failure epoch through the equivalence $[\tilde{t}_n \leq T] \iff [n(T) \geq n]$. The mean and variance of $\tilde{n}(T)$ are, of course, $\phi G(T) + \omega T$. Moments of the lifetimes $\{\tilde{x}_j\}$, on the other hand, have no simple form in the general case; however, the concept of *local mean life* at t , introduced in [4], and defined as:

$$(8.3) \quad \mu(t) = 1/\lambda(t) = [\phi g(t) + \omega]^{-1},$$

has the interpretation of mean duration until the next failure if no further change occurs in $g(t)$ (or if it is slowly varying). Of course, if $g(t) \rightarrow 0$ as $t \rightarrow \infty$, successive lifetimes approach independent, exponentially-distributed random variables with mean life $\mu(\infty) = \omega^{-1}$.

Models of type II with $h(x) \equiv 1$, on the other hand, give rise to independent, exponentially-distributed random lifetimes for every interval; the j^{th} lifetime has mean:

$$E\{\tilde{x}_j \mid t_{j-1}; \phi, \omega\} = \mu(t_{j-1}) = [\phi g(t_{j-1}) + \omega]^{-1}.$$

Note that this point process is *not*, however, a Poisson process, since the intensity parameter jumps to its new value only at a failure epoch, and not continuously over time. Hence, there is no simple form for the distribution of $\tilde{n}(t)$.

9. BAYESIAN ESTIMATION

Because most product development programs involve systems which are evolutions of previous ones, it makes sense to quantify this engineering experience and "know-how" by using Bayesian estimation procedures. There appear to be three major ways in which this approach might be applied in reliability growth.

A. Full Bayesian Estimation

In a full Bayesian analysis, the likelihood (5.6) (or an extended version developed in Section 7) is used together with a joint *prior density* $p(\phi, \omega)$ in the usual Bayes' formula to find the *posterior joint density*:

$$(9.1) \quad p(\phi, \omega \mid \mathcal{D}) = \frac{L(\phi, \omega \mid \mathcal{D})p(\phi, \omega)}{\text{normalization } (\mathcal{D})} .$$

There is probably little to be gained by attempting an analytic solution to this equation through the use of natural conjugate priors. Even if $p(\phi, \omega)$ factored into independent gamma densities on ϕ and ω , it can be seen from (5.8) that the posterior-to-date density would have n terms, corresponding to different priors $\phi^j \omega^{n-j}$ ($j = 0, 1, \dots, n$).

If, in addition to ϕ and ω , a hidden parameter in g or h must be estimated, this leads to a very non-exponential-type likelihood, for which no natural conjugate prior exists.

Thus, in general, it seems that useful full Bayesian solutions of (9.1) would require 2- or 3-dimensional numerical integration.

B. Multinormal Approximation

If the prior and the likelihood are appropriately bell-shaped functions of ω and ϕ , then one can make multinormal approximations

of the densities, and then carry out an exact calculation using (9.1). The parameters for the likelihood would have to be calculated numerically using the formulas in Section 6; the multinormal prior would require eliciting the prior means and covariances of ω and ϕ from engineering personnel, or examining failure rate records of related development programs. At the very least, this requires knowing how much data and prior confidence is needed to obtain unimodal densities of the correct form.

C. Conditional Likelihood

The estimation problem is greatly simplified if we can assume that only one parameter is imprecisely known. A possible situation in which this might occur is where the *shape of the learning curve* and the *initial failure rate* $f(0,0 \mid 0) = \alpha$ are well-known as testing begins. Assuming that the prototypical failure functions are normalized so that $h(0) = g(0 \mid 0) = 1$, we set $\alpha = \phi + \omega$, and use the one-dimensional likelihood and prior in:

$$(9.2) \quad p(\omega \mid \mathcal{D}, \alpha) = \frac{L(\omega \mid \mathcal{D}, \alpha) p(\omega \mid \alpha)}{\text{normalization}}$$

$$(9.3) \quad L(\omega \mid \mathcal{D}, \alpha) = \prod_{i=1}^S \prod_{j=1}^{n_i} (\alpha g_{ij} + \omega(h_{ij} - g_{ij})) \exp \left\{ -\alpha \sum_{i=1}^S G_i - \omega \sum_{i=1}^S (H_i - G_i) \right\}.$$

Thus, attention is shifted to estimation of the *ultimate failure rate*, ω , which is usually the performance parameter of greatest interest.

It shall be remembered though, that α (and γ) are assumed known. For instance, Figure 3 shows a typical prior density $p(\omega \mid \alpha)$, with experience indicating a most likely improvement of about 40% from the initial failure rate α , but with allowances being made for the possibility of other outcomes, even $\omega > \alpha$ (reliability loss). This prior

would then be multiplied by $L(\omega \mid \mathcal{D}, \alpha)$ in the usual way to give the (unnormalized) posterior $p(\omega \mid \mathcal{D}, \alpha)$. The point estimate of ω would then depend upon the loss function selected.

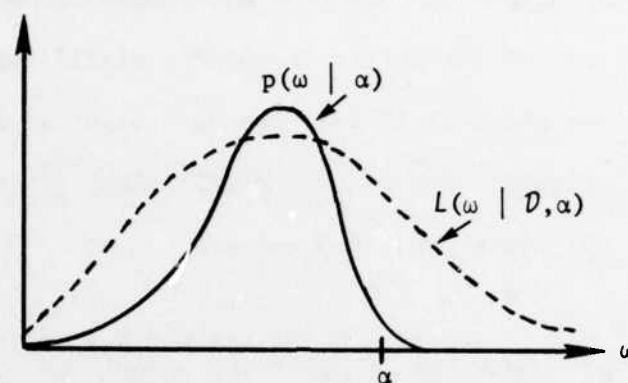


FIGURE 3. Illustration of Bayes' Law for Conditional Likelihood Approach.

10. NUMERICAL TRIALS - EXPONENTIAL GROWTH FUNCTION

A. Random Variates

To illustrate the numerical difficulties associated with reliability growth models, a series of random variates were generated using constant hazard rate, an exponential growth function (assuming Model I is operating), and a generous set of parameters in which a significant decrease in failure rate occurs after about 15-20 failures have occurred in each realization.

More specifically, pseudo random lifetimes $\{x_{ij} ; i = 1, 2, \dots, 32 ; j = 1, 2, \dots, 80\}$ were generated using:

$$\begin{aligned} h(x) &= 1 & \alpha &= 1.0 ; \omega = 0.2 \quad (\emptyset = 0.8) \\ g(t) &= e^{-\gamma t} & \gamma &= 0.05 . \end{aligned}$$

This gives a time-varying Poisson process in which the failure rate decreases from an initial rate of 1.0 towards an ultimate rate of 0.2, with a time constant of $\gamma^{-1} = 20$ time units.

Figure 4 shows the resulting 32 point process realizations over the range $0 \leq T \leq 40$, corresponding to zero to two time constants, which is probably the range of practical interest, as only 13.5% of the growth is unobserved at $T = 40$. (Because of quantization in the printing of Figure 4, failures that are closer together than 0.5 time unit show as only one event; for example, in realization #1, the failures near $T = 2.5$, 5.0, and 14.5 are actually double events.) Even with the high variability of the Poisson, the strong reduction in failure rate with time can be visually inferred.

The average number of events in one realization follows the law:

$$M(t) = E\{\tilde{n}(t)\} = 0.2t + 16(1 - e^{-0.05t}) .$$

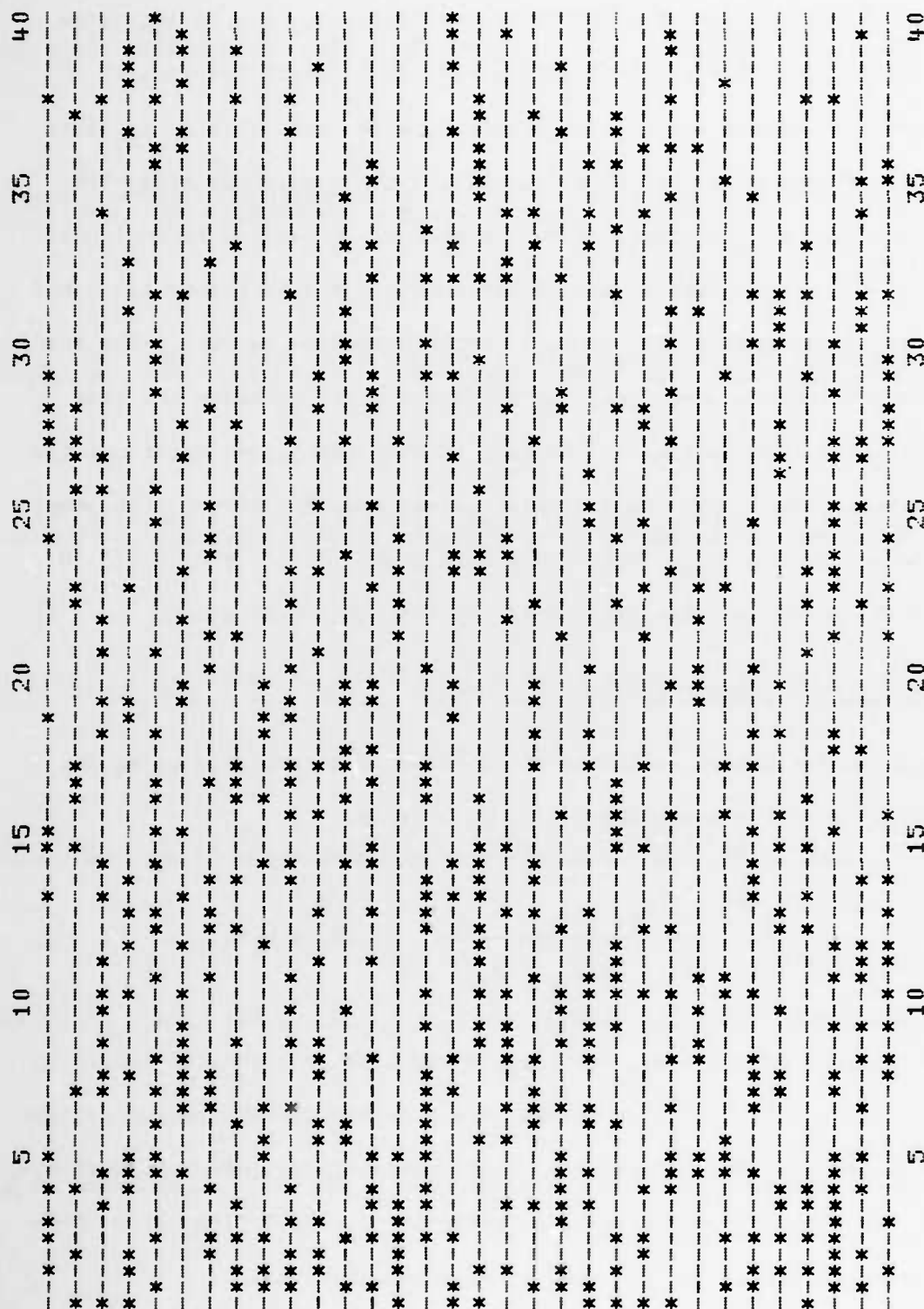


FIGURE 4. 32 point-process realizations used in numerical trials.
 Failure rate = $0.2 + 0.8 \exp(-.05t)$. (Epochs closer
 together than 0.5 time units shown as single event.)

Figure 5 is intended to show how this form is better and better approximated by $\sum n_i(T)/S$, as the number of realizations, S , is increased. First of all, for $S = 1$ (realization #1), this sample function is just the usual unit-step counting function, with jumps at the t_{1j} ; it starts out rather well, but then lags behind the mean, beginning with the rather long 8^{th} lifetime. When the second realization is included, making $S = 2$, the jumps in the sample function are of height $1/2$, and occur at failure epochs $\{t_{1j}; t_{2j}\}$ of both systems on test. The resulting curve is somewhat smoother, but of course still contains the bias of the results from system #1. Keeping in mind the jagged nature of the actual sample functions, the remaining curves show how closely the sample function approaches $M(t)$ for $S = 4, 8, 16$, and 32 , for $T = 0(5)40$. The $S = 32$ curve deviates from $M(t)$ by less than 0.26 units.

B. Maximum-Likelihood Estimates

Maximum-likelihood estimates of ω were then evaluated using the (9.3) version of the equations of (7.3), that is:

$$(10.1) \quad \sum_{i=1}^S \sum_{j=1}^{n_i(T)} \frac{h_{ij} - g_{ij}}{\alpha g_{ij} + \hat{\omega}(h_{ij} - g_{ij})} = \sum_{i=1}^S (H_i - G_i) .$$

In our case, of course, $h_{ij} = 1$, $g_{ij} = \exp(-\gamma t_{ij})$, and the RHS is just $S[T - G(T)]$. These values were obtained through an interactive numerical search over successively finer values of ω ; when the maximizing $\hat{\omega}$ was obtained to within 0.01, parabolic interpolation using three neighboring values of ω was used to estimate the final $\hat{\omega}$.

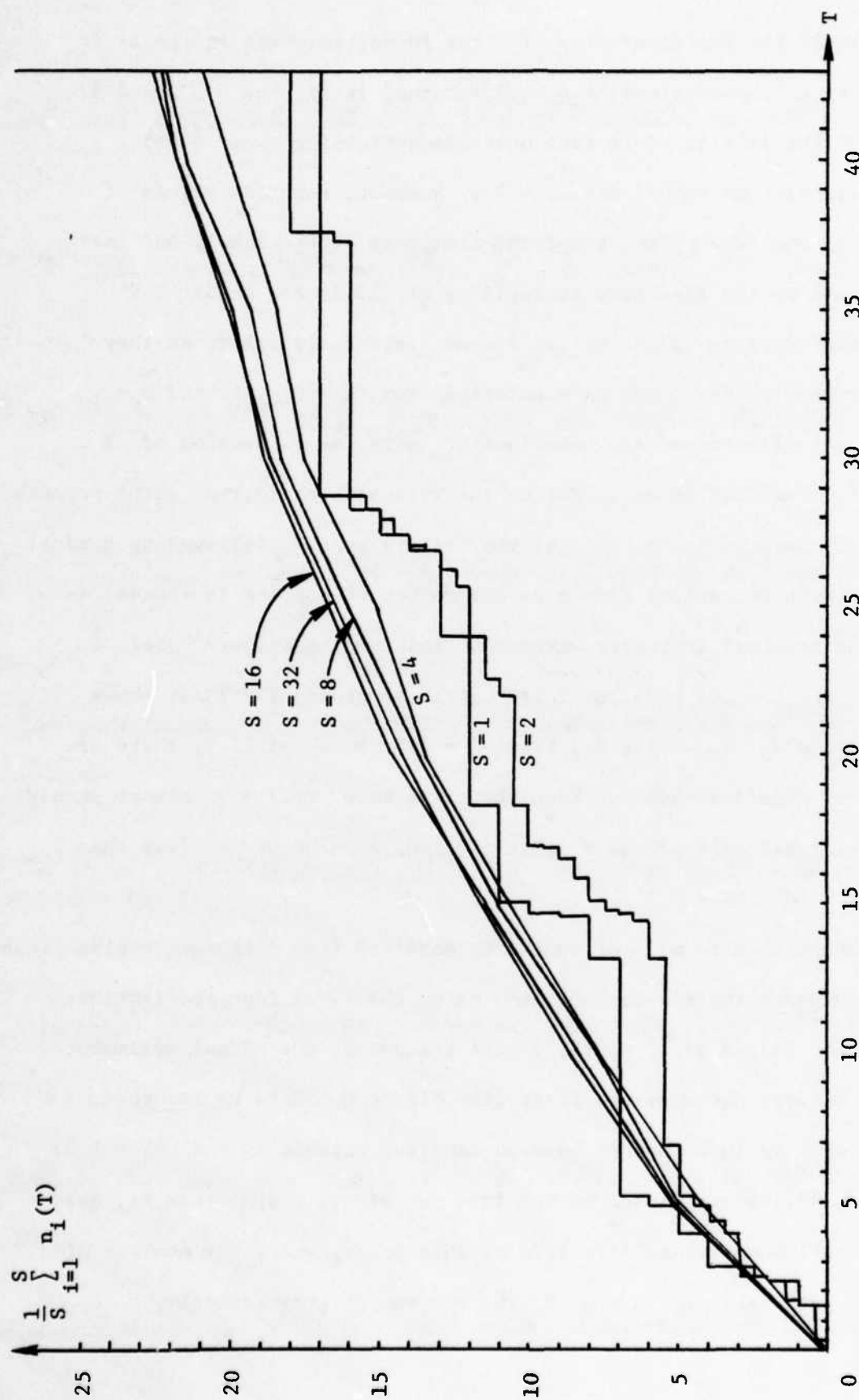


FIGURE 5. Sample functions for $M(T)$ from simulations with $S = 1, 2, 4, 8, 16, 32$ systems. Only selected values of T shown for larger values of S . (Straight lines added for continuity.)

For small testing intervals, T , the formal solution of (10.1) is often negative; these values have been retained in Figures 6, 7, and 8, even though the likelihood is then monotone decreasing over $(0, \infty)$, and theoretically we should set $\hat{\omega} = 0$. However, negative values of $\hat{\omega}$ give us some idea of the resulting steepness of the likelihood near zero, as well as the algebraic instability of (10.1) for small T . However, the negative values of $\hat{\omega}$ are not especially exact, as they were determined by parabolic extrapolation from $\omega = 0, .01$, and 0.2 .

Figure 6 illustrates the behaviour of $\hat{\omega}(T)$ as a function of T for $S = 1$ (realization #1). Notice the typical behaviour of point process estimators; jumps in value occur at the failure epochs, followed by gradual decreases (with decreasing slopes as the number of samples increases) as the testing interval increases without an intervening failure. For T from 0 to t_{11} , $\hat{\omega}$ is undefined; from t_{11} to about $T = 7$, it shows *reliability decay* ($\hat{\omega} > 1$), and from $T = 7.5$ to about 27.5, there are intervals of negative values. Even though we know $\hat{\omega}(T) \rightarrow \omega$ almost surely as $T \rightarrow \infty$, realization #1 is still estimating a value of ω less than 0.01 at $T = 50$ ($\gamma T = 2.5$).

Different results will of course be obtained from different realizations. Figure 7 compares the results obtained using the first four realizations. Only selected values of $T = 5(5)50$ are evaluated; the actual estimator behaviour between these points looks like Figure 6. Here we see the large variability in behaviour between samples, especially for $\gamma T < 1.0$. Samples #2, #3, #4 are closer to the true ω at $\gamma T = 2.5$ than #1, but there is still some variability left at this point, where the numbers of samples in each realization are 19, 19, 28, and 26, respectively.

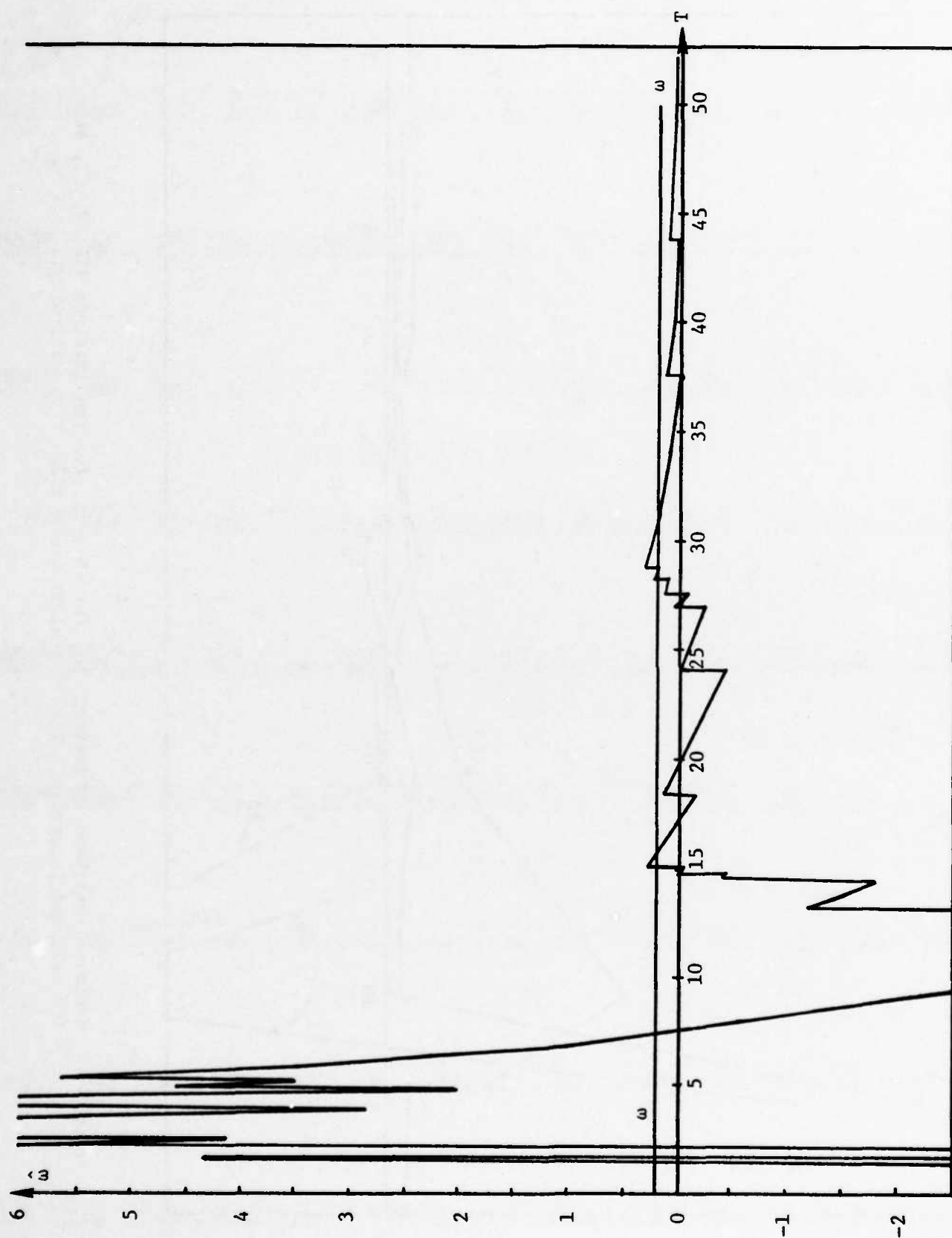


FIGURE 6. Maximum-likelihood estimator, $\hat{\omega}$, for $S = 1$, using realization #1, versus testing interval, T . (See text regarding negative values.)

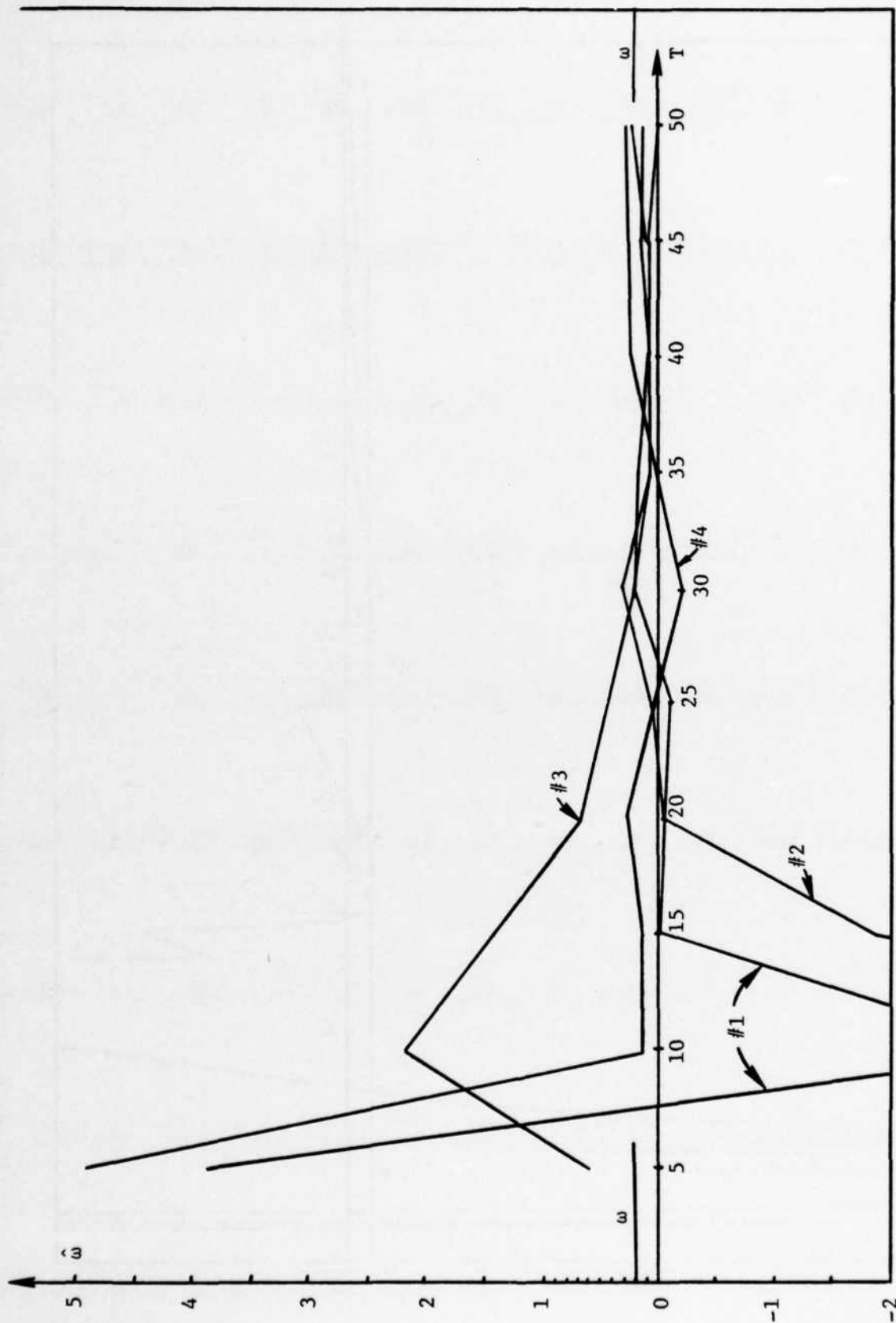


FIGURE 7. Maximum-likelihood estimator, $\hat{\omega}$, for $S = 1$, for realizations #1, #2, #3, #4, for selected intervals, T . (Straight lines added for continuity.)

Clearly, the estimator will be more stable if we increase S , the number of systems on test. Figure 8 shows $\hat{\omega}(T)$ versus $T = 5(5)50$ for $S = 2, 4, 8, 16$, and 32 . $S = 2$ still has a large region of negative results (because $\hat{\omega}$ of both realizations #1 and #2 in Figure 7 are negative), but $S = 4$ and higher are relatively stable after $T = 20$ ($\gamma T = 1.0$). Remember that the complete behaviour between selected T should look more like Figure 6, but that with increasing S , the jumps will be smaller, and will occur more often. Note also that the vertical scale of Figure 8 is finer than that of Figures 6 and 7, so that the relative stability is better than it appears. The numbers of samples at $T = 50$ are 38, 92, 195, 395, and 787 for $S = 2, 4, 8, 16$, and 32 , respectively.

C. Data Information Estimates

For the cases in which an internal maximum occurs in the likelihood, it is convenient to have a measure of how sharp this maximum is. A convenient choice is the (*Fisher*) *Information*, defined as:

$$(10.2) \quad I = E \left[\frac{\partial^2 \ln L(\omega | \tilde{D})}{\partial \omega^2} \mid \hat{\omega} \right].$$

If I as a function of ω is approximated by a normal curve about $\hat{\omega}$, then I is approximately the *precision* (inverse variance) of the approximation. Estimates of the information, \hat{I} , were obtained during the course of the computations of $\hat{\omega}$; if the formal solution of (10.1) gave negative values, \hat{I} was computed at these maxima, but, of course, the information then has no useful physical interpretation.

Figure 9 shows the values of \hat{I} obtained at selected values of $T = 5(5)50$ for different values of S . Points with negative $\hat{\omega}$ are

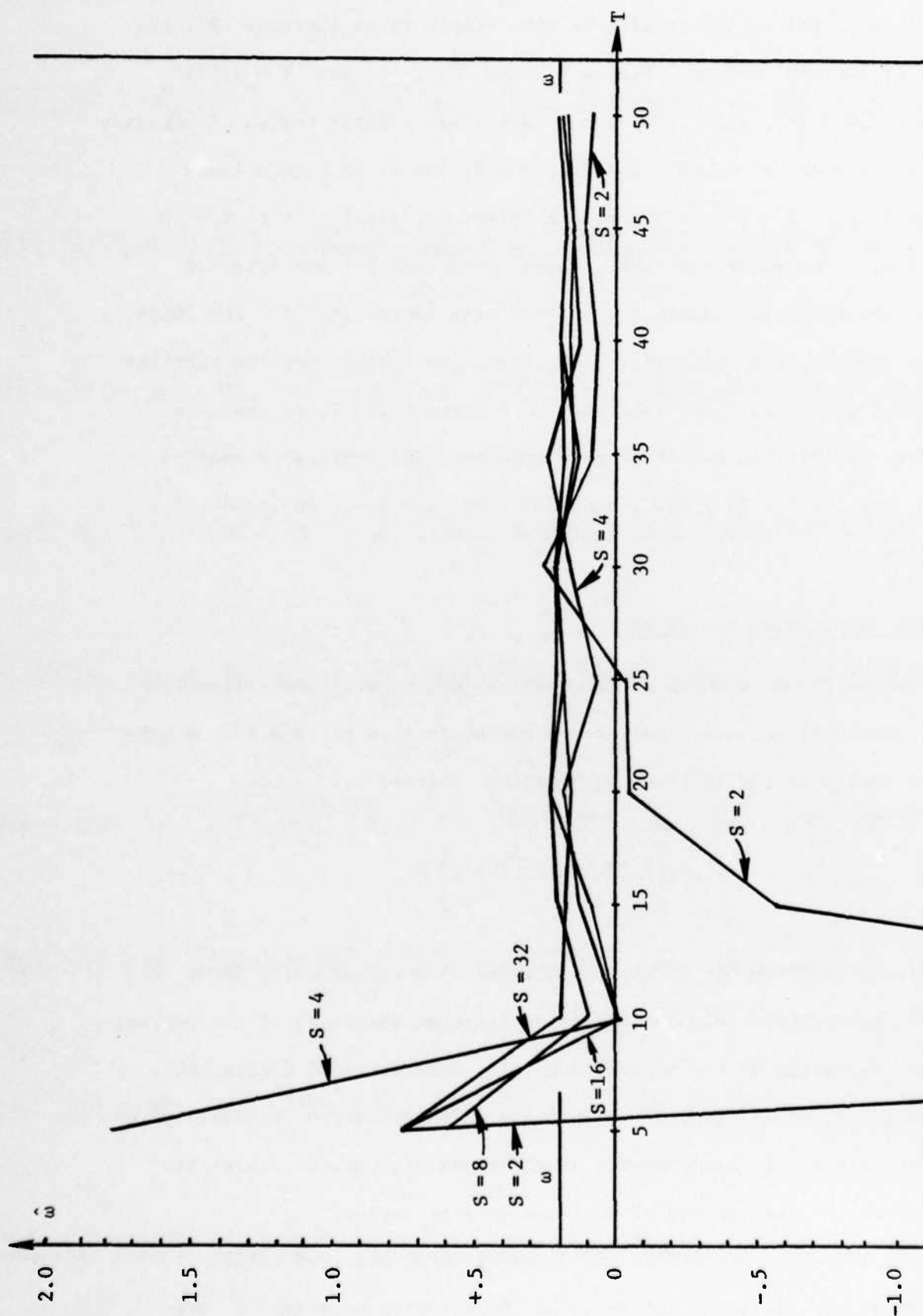


FIGURE 8. Maximum-likelihood estimator, $\hat{\omega}$, for $S = 2, 4, 8$, and 16 systems, for selected testing intervals, T . (Straight lines added for continuity.)

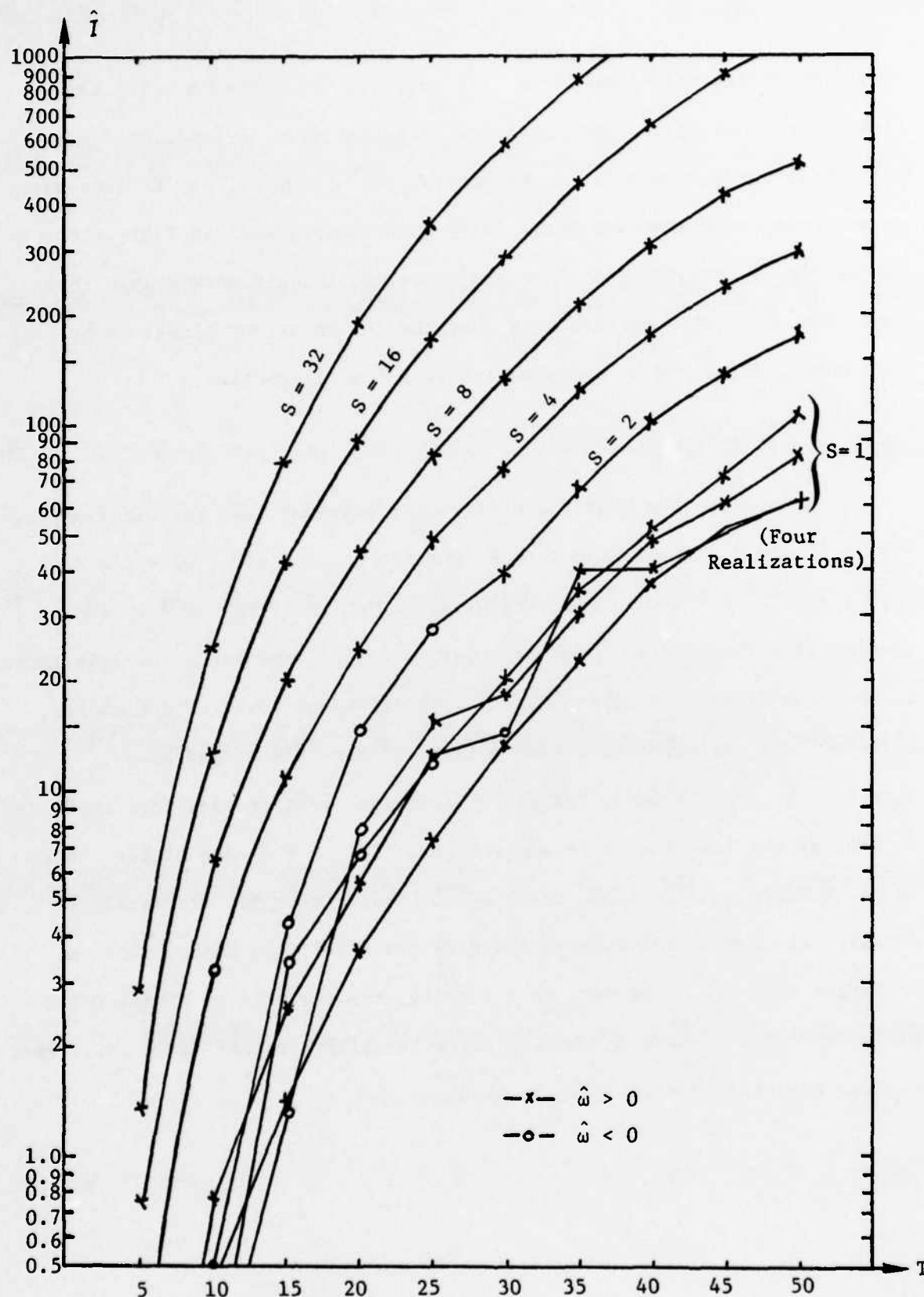


FIGURE 9. Data information estimator, \hat{I} , for four different realizations of $S = 1$, and for $S = 2, 4, 8, 16$, and 32 , for selected testing intervals, T . (Straight lines added for continuity.)

indicated differently than those with positive $\hat{\omega}$. For $S = 1$, the first four realizations lead to quite different results (and, of course, the actual curve versus T is discontinuous). However, as S increases, we may expect the sampling variability to decrease, and, in fact, the curves for S greater than four are quite smooth, and more stable than those for $\hat{\omega}$. Note particularly that the curves appear almost to be the same curve, displaced by the same amount as S is doubled.

D. Data Likelihoods

To give some idea what the different precisions mean for our example, Figure 10 shows the likelihood as a function of ω , $p(\mathcal{D} \mid \omega) = L(\omega \mid \mathcal{D})$, for $T = 20$ ($\gamma T = 1.0$), and various values of S . For $S = 1$ and 2, the formal solution for $\hat{\omega}$ gives negative values, and hence the likelihood is monotone decreasing over $(0, \infty)$; with different realizations chosen, different results might have been obtained. For $S = 4$ and above, internal $\hat{\omega}$ are obtained, and the information \hat{I} increases from about 24 to 196, as the likelihoods become sharper. For $S = 2$ and higher, there is practically no likelihood that ω is greater than 1.0 (reliability decay); for $S = 8$ and higher, there is practically no chance that ω is larger than 0.7. However, in all cases, the instability of the point estimate $\hat{\omega}$ is clearly shown, and there is still a rather large confidence region, even for $S = 32$ (455 samples).

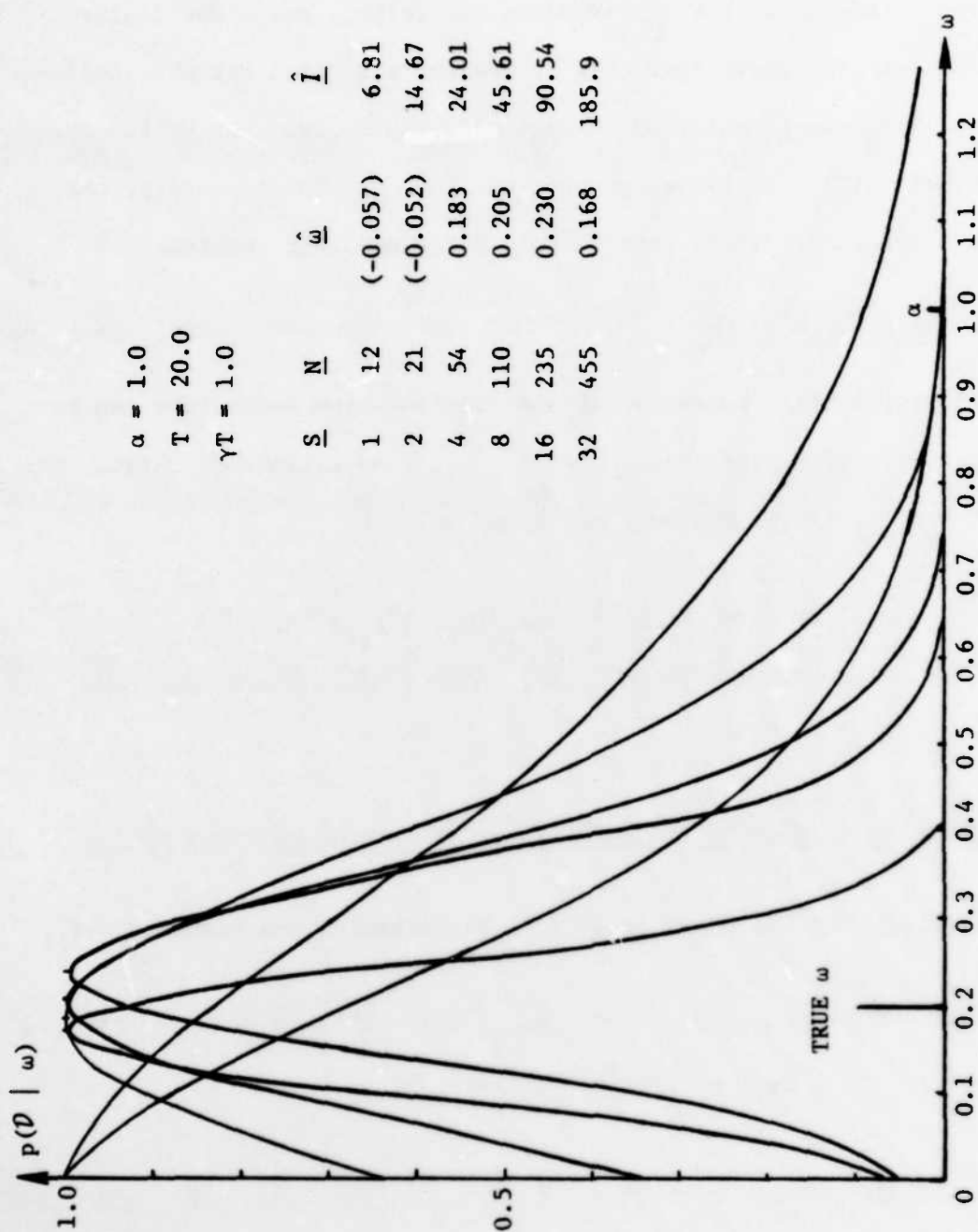


FIGURE 10. Data likelihoods, $p(D | \omega) = L(\omega | D)$ versus ω , for $S = 1, 2, 4, 8, 16$, and 32 systems on test, and testing interval $T = 1.0$, showing total number of samples, N , maximum likelihood estimator, $\hat{\omega}$, and data information estimator, \hat{I} . (Likelihoods normalized so that maximum value = 1.0.)

11. ANALYTIC RESULTS ON ESTIMATORS

Although the numerical results of the last section were enlightening, they were obtained for particular parameter values, and a particular choice of learning curve form. It is desirable to also ask what additional results can be obtained analytically, apart from the well-known result that $\hat{\omega}(T) \rightarrow \omega$ almost surely, as $T \rightarrow \infty$. Unfortunately, the complicated form of (10.1) permits only a few explicit results.

A. Initial Behaviour for $S = 1$

The poor initial behaviour of the single-system estimators can be explained by noting first that, for $T < t_{11}$, no estimators exist. For $t_{11} \leq T < t_{12}$, (10.1) has only one term, so that:

$$(11.1) \quad \hat{\omega}(T) = (H_1 - G_1)^{-1} - \alpha g_{11} (h_{11} - g_{11})^{-1},$$

and

$$(11.2) \quad \hat{I}(T) = (H_1 - G_1)^2.$$

In particular, for the forms assumed in the example, and assuming γt_{11} is small, we find:

$$(11.1') \quad \hat{\omega}(T) \approx (2/\gamma T^2) - \alpha(1/\gamma t_{11}),$$

$$(11.2') \quad \hat{I}(T) \approx \gamma^2 T^4/4.$$

This means that the initial MLE estimator, $\hat{\omega}(t_{11})$, has the following sensitivity to the first lifetime:

$$\begin{aligned}
 & \text{if } \alpha t_{11} > 2, & \text{then } \hat{\omega}(t_{11}) < 0; \\
 (11.3) \quad & \text{if } 2(1 - (2\gamma/\alpha)) < \alpha t_{11} < 2, & \text{then } 0 < \hat{\omega}(t_{11}) < \alpha; \\
 & \text{and if } \alpha t_{11} < 2(1 - (2\gamma/\alpha)), & \text{then } \hat{\omega}(t_{11}) > \alpha
 \end{aligned}$$

(assuming that γ is small compared to α). Since only the middle estimate corresponds to reliability growth, this means that the initial estimates of ω are likely to be pathological ones, since the range of allowed t_{11} is small ($1.9 < t_{11} < 2.0$ in our example).

The estimator of the precision, on the other hand, is a smooth function of T , and indicates how quickly precision builds up after $T > t_{11}$, until the next discontinuity at $T = t_{12}$.

B. Initial Behavior for Arbitrary S

The above discussion can be extended to an arbitrary number of systems, provided we limit our discussion to values of T between the first failure to occur,

$$t_{*1} = \min(t_{11}, t_{21}, \dots, t_{S1}),$$

and the second to occur. This will occur, on the average, earlier with increasing S , so the approximations are even better. (11.1) and (11.2) are first modified by replacing $H_1 - G_1$ by the sums $\sum (H_i - G_i)$ and h_{11} and g_{11} become h_{*1} and g_{*1} , in an obvious notation. Then (11.1) and (11.2) become:

$$(11.1'') \quad \hat{\omega}(T) \approx (2/S\gamma T^2) - \alpha(1/\gamma t_{*1}),$$

$$(11.2'') \quad \hat{I}(T) \approx S^2 \gamma^2 T^4 / 4.$$

Not only is $\alpha t_{*1} < (2/S)$ the new requirement that $\hat{\omega}(t_{*1})$ be positive, the other bound becomes: $(2/S)(1 - (2\gamma/\alpha S))$, thus *widening*, relatively, the gap in which t_{*1} must fall to give an initial estimate corresponding to reliability growth! The quadratic dependence of \hat{I} on S is also reassuring.

Unfortunately, discussion of the next interval of T , beyond the second failure, already leads to quadratic expressions for $\hat{\omega}(T)$.

C. Average Behavior for γT Small and S Large

Before giving the results of this section, let us recall what kind of results we would get if we were estimating the failure rate λ of a *stationary Poisson process*, with S systems on test for T time periods. The maximum likelihood estimator of λ is just $\hat{\lambda}(T) = N(T)/ST$, and it is immediate that this estimator is unbiased for all T , since $E\{\hat{\lambda}(T)\} = SM(T)/ST = \lambda$ for all T . Furthermore, $E\{\hat{I}(T)\} = SM(T) = S\lambda T$, so that the average amount of precision in any experiment is proportional to ST , the total-time-on-test.

To obtain similar results using (10.1), using $h(t) = 1$ but arbitrary $g(t)$, we first of all set T small enough so that the probability of more than one failure per system in $(0, T]$ is negligible compared to the probability of zero or one failure. If

$$(11.4) \quad g(t) = 1 - \gamma_1 t + \gamma_2 t^2/2 - \gamma_3 t^3/6 + \dots \quad (t \rightarrow 0),$$

this means that $\gamma_1 T \ll 1$, and the other γ_i do not increase explosively (which is certainly true if $g(t)$ is decreasingly slowly and monotonically). For $g(t) = \exp(-\gamma t)$, the $\gamma_i = \gamma^i$, and we require $\gamma T \ll 1$.

Letting $t_1, t_2, \dots, t_{N(t)}$ denote the random number of (first) failures which would be observed in such a case, we rewrite (10.1) as:

$$(11.5) \quad T - G(T) = \frac{1}{S} \sum_{k=1}^{N(t)} \frac{1 - g(t_k)}{\alpha g(t_k) + \hat{\omega}(1 - g(t_k))} .$$

Now the \tilde{t}_k are independent, identically distributed random variables with the same distribution as the $\{\tilde{t}_{i1} \mid \tilde{t}_{i1} \leq T\}$, that is, they have a normalized density with failure rate $g(t)$ that is approximated by (11.4). With a little bit of labor, we find the first two moments of these first failure epochs to be:

$$(11.6) \quad E\{\tilde{t}_k\} \approx \frac{T}{2} - \frac{\alpha^2 + (\alpha - \omega)\gamma_1}{12\alpha} T^2 + O(T^3)$$

$$E\{\tilde{t}_k^2\} \approx \frac{T^2}{3} + O(T^3) .$$

Now let the number of systems on test get very large; then (11.5) becomes:

$$T - G(T) = E \left\{ \frac{1 - g(\tilde{t}_k)}{\alpha g(\tilde{t}_k) + \omega(1 - g(\tilde{t}_k))} \right\} \text{ almost surely as } S \rightarrow \infty .$$

We now expand both sides of this equation in powers of T , using (11.4) and (11.6), assuming that $E\{\hat{\omega}(T)\} \approx \omega_0 + \omega_1 T + O(T^2)$. After some labor, we find that the first and second powers of T are trivially satisfied, but that from the third power we get:

$$(11.7) \quad E\{\hat{\omega}(T)\} \approx \omega - \frac{\alpha^2}{4\gamma_1} + O(T) \quad (\gamma_1 T \ll 1) .$$

This means that the maximum likelihood estimator is *inconsistent* at small values of T --in fact, for the parameters of the numerical trials, $\hat{\omega}(T)$ approaches $\omega = 5.0$! This gives us additional evidence as to the unsuitability of the maximum likelihood estimator.

Using similar techniques, we can estimate the average information in the data for small T as:

$$(11.8) \quad E\{\hat{I}(T)\} \approx S \frac{\gamma_1^2 T^3}{3\alpha} + O(T^4) \quad (\gamma_1 T \ll 1) .$$

Since the mean number of samples is of the order of $S\alpha T$, this shows that the precision of the estimate is increasing strongly with increasing time-on-test T , not just proportionally to total-time-on-test ST .

For the data in the numerical trials, this estimate of precision is $ST^3/1200$. Table I shows that this estimate is, in fact, rather good for $\gamma T \leq 1.0$, and $S \geq 4$, thus explaining the regularity observed in Figure 9. Remember, however, that this refers to the second derivative of the \ln -likelihood at a possibly negative $\hat{\omega}$.

	S = 4	S = 8	S = 16	S = 32
T = 5	0.33 0.42	0.76 0.83	1.38 1.67	2.86 3.33
T = 10	3.33 3.33	6.46 6.66	12.9 13.3	24.4 26.7
T = 20	10.5 11.3	20.3 22.5	43.3 45	79.9 80.0
T = 30	24.0 26.7	45.6 53.5	90.5 107	186 213

TABLE I. Data Information, \hat{I} , for Different S and T , as Estimated from Numerical Trials (upper numbers) and from Approximation (11.8) (lower numbers).

D. Bayesian Estimators

This is perhaps a convenient point to re-emphasize that none of the above difficulties occur when using a Bayesian analysis. Not only does (9.2) give the posterior-to-experiment distribution of $\tilde{\omega}$ for any testing protocol, it follows then that, given a risk function, we can make point estimates of the parameters, or, if desired, can specify Bayesian tolerance regions for $p(\omega \mid \mathcal{D})$. If either the prior or the likelihood is approximately normal, we can approximate the posterior-to-experiment mean, $E(\tilde{\omega} \mid \mathcal{D})$, by a precision-weighted "credibility" [12] mixture of the prior mean, $E(\tilde{\omega})$, and the MLE $\tilde{\omega}$; the precision of this estimate is just the sum of the prior precision and the data information. And so on. There are no conceptual difficulties with estimating more parameters, although, as indicated above, numerical integration would be necessary in most cases.

12. CONCLUSIONS

Before summarizing the results of this study, we recall the basic assumptions behind our learning-curve reliability growth model: only the past epochs of failure may be used to estimate process parameters (although the modes of failure may be used in making the actual improvements); the date of the most recent failure of a system is a surrogate for all previous history of failures and modifications of that system, that is, the system state is Markovian at these epochs and depends only upon clock time and local age; and, the age and global time failure rate effects are separable, as in (2.2).

Given this framework, we believe the following conclusions can be made:

- (1) A complete learning curve reliability growth model requires at least two parameters and two prototypical failure rate functions to describe the desired phenomenon;
- (2) If the learning-curve effect is continuous and slow, the exact form of the prototypical growth function g is probably not too important; there is also not much difference between the "as operated" and "as produced" specializations introduced in Section 3;
- (3) However, use of the Duane learning curve, $g(t) = Kt^{\gamma-1}$, leads to technical difficulties in the problems of interest. An exponential learning curve, $g(t) = e^{-\gamma t}$, avoids these difficulties, and is, moreover, suggested by certain defect removal models of software reliability growth;
- (4) In any case, the mathematical problems of finding the maximum-likelihood estimates of the parameters in (6.1), (6.2), (10.1),

- or the data information (matrix) measure (6.3), (6.4), (6.5), (10.2), are similar for the different models and forms;
- (5) Limited computational experience in estimating the ultimate failure rate, ω , under favourable assumptions shows that the maximum-likelihood estimator, $\hat{\omega}$, is very unstable during the early part of the learning curve ($\gamma_1 T < 1$), especially for a small number S of systems on test. In fact, for $\gamma_1 T$ small and S large, $\hat{\omega}$ is inconsistent, usually much less than the true ω , and negative;
 - (6) The estimate of the data information measure, on the other hand, is remarkably stable. (11.8) is apparently valid as a scaling law for $\gamma_1 T \leq 1$ and $S \geq 4$, and can be used to estimate the precision of relatively non-informative testing protocols. In any case, it is always better to increase T than to increase S in this region.

Thus, the estimation of ultimate performance using classical estimation methods presents fundamentally difficult problems, unless testing periods are extended well into the learning curve, and a reasonable number of systems are placed on test. The problem of simultaneous estimation of several parameters will likely introduce additional inconsistencies and instabilities.

Bayesian estimation procedures, on the other hand, use data from *any* testing protocol in a consistent and "friendly" manner, with the posterior-to-testing density of ω given directly from the likelihood and the prior. Point estimates or Bayesian tolerance regions can then be obtained through numerical integration, or through approximations, if either the prior or the likelihood has high precision. We expect to see further development

of reliability growth procedures to place more emphasis upon the use of engineering experience obtained from related development testing programs.

REFERENCES

- [1] Bain, Lee J. and Max Engelhardt, "Inferences on the Parameters and Current System Reliability for a Time Truncated Weibull Process," Technometrics, Vol. 22, No. 3, pp. 421-426 (August 1980).
- [2] Barlow, Richard E. and Ernest M. Scheuer, "Reliability Growth During a Development Testing Program," Technometrics, Vol. 8, pp. 53-60 (1966).
- [3] Barr, Donald R., "A Class of General Reliability Growth Prediction Models," Operations Research, Vol. 18, No. 1, pp. 52-65 (January-February 1970).
- [4] Braun, Henry I. and J. Macon Paine, "A Comparative Study of Models for Reliability Growth," Technical Report No. 126, Series 2, Department of Statistics, Princeton University (1977).
- [5] Braun, Henry I., Nathaniel Schenker and J. Macon Paine, "A Comparative Study of Models for Reliability Growth," Department of Statistics, Princeton University (December 1980).
- [6] Corcoran, W. J., H. Weingarten and P. W. Zehna, "Estimating Reliability After Corrective Action," Management Science, Vol. 10, No. 4, pp. 786-795 (July 1964).
- [7] Crow, Larry H., "Reliability Growth Modelling," Technical Report No. 55, U.S. Army Materiel Systems Analysis Agency, Aberdeen Proving Ground, MD (August 1972).
- [8] Crow, Larry H., "Reliability Analysis for Complex, Repairable Systems," in RELIABILITY AND BIOMETRY, F. Proschan and R. J. Serfling (eds.), SIAM, Philadelphia, PA, pp. 379-410 (1974).
- [9] Crow, Larry H., "Confidence Interval Procedures for the Weibull Process with Applications to Reliability Growth," Technometrics, Vol. 24, No. 1, pp. 67-72 (February 1982).
- [10] Department of Defense, RELIABILITY GROWTH MANAGEMENT, MIL-HDBK-189, Washington, DC, 20301 (February 13, 1981).
- [11] Dynamics Group, "Reliability Aspects of Microprocessor Systems," ST-25358, British Aerospace (July 1981).
- [12] Jewell, William S., "Bayesian Life Testing Using the Total Q on Test," ORC 76-3, Operations Research Center, University of California, Berkeley (January 1976). Also in THE THEORY AND APPLICATIONS OF RELIABILITY, C. P. Tsokos and I. N. Shimi (eds.), Academic Press, New York, pp. 49-66 (1977).
- [13] Jewell, William S., "Bayesians Learn While Waiting," ORC 77-19, Operations Research Center, University of California, Berkeley (July 1977). Also in The Astin Bulletin, Vol. 10, Part 2, pp. 163-172 (March 1979).

- [14] Jewell, William S., "'Reliability Growth' as an Artifact of Renewal Testing," ORC 78-9, Operations Research Center, University of California, Berkeley (June 1978).
- [15] Jewell, William S., "Stochastically-Ordered Parameters in Bayesian Prediction," ORC 79-12, Operations Research Center, University of California, Berkeley (October 1979).
- [16] Jewell, William S., "A Curious Renewal Process Average," Stochastic Processes and their Applications, Vol. 11, pp. 293-295 (1981).
- [17] Lee, Larry, "Comparing Rates of Several Independent Weibull Processes," Technometrics, Vol. 22, No. 3, pp. 427-430 (August 1980).
- [18] Littlewood, B. and J. Verrall, "A Bayesian Reliability Growth Model for Computer Software," Applied Statistics (J. Roy. Statist. Soc., Series C), Vol. 22, No. 3, pp. 332-346 (1973).
- [19] Musa, John D., "A Theory of Software Reliability and its Application," IEEE Transactions on Software Engineering, Vol. SE-1, pp. 312-327 (September 1975).
- [20] Singpurwalla, N., "An Adaptive Bayesian Scheme for Estimating Reliability Growth Under Exponential Failure Times," ORC 79-17, Operations Research Center, University of California, Berkeley (December 1979).
- [21] Smith, A. F. M., "A Bayesian Note on Reliability During a Development Program," IEEE Transactions on Reliability, Vol. R-26, pp. 346-347 (1977).

END

FILMED

9-83

DTIC